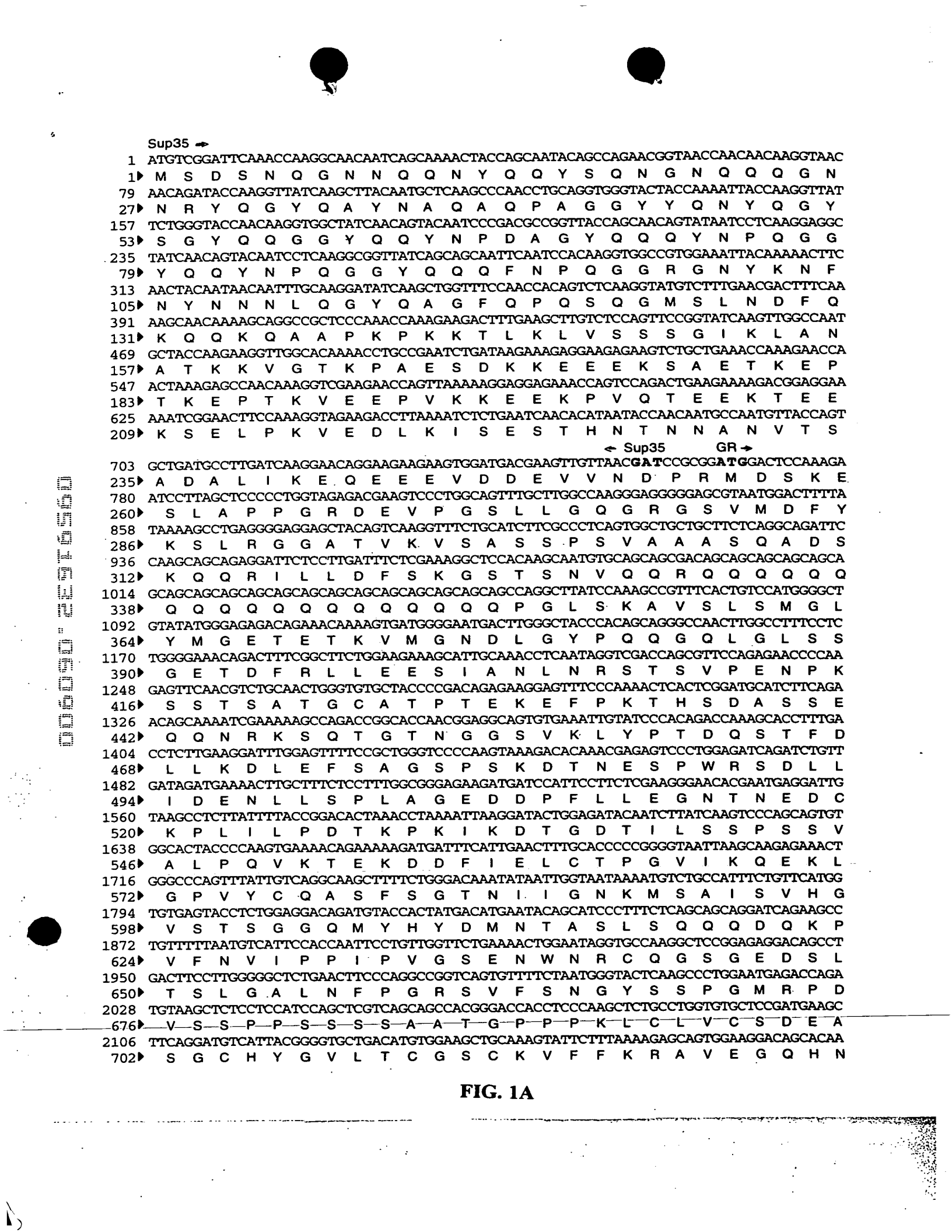
Sup35 →
1 ATGTCGGATTCAAACCAAGGCAACAATCAGCAAAACTACCAGCAATACAGCCAGAACGGTAACCAACAACAAGGTAAC
1 M S D S N Q G N N Q Q N Y Q Q Y S Q N G N Q Q Q G N
79 AACAGATACCAAGGTTATCAAGCTTACAATGCTCAAGCCCAACCTGCAGGTGGGTACTACCAAAATTACCAAGGTTAT
27 N R Y Q G Y Q A Y N A Q A Q P A G G Y Y Q N Y Q G Y
157 TCTGGGTACCAACAAGGTGGCTATCAACAGTACAATCCCGACGCCGGTTACCAGCAACAGTATAATCCTCAAGGAGGC
53 S G Y Q Q G G Y Q Q Y N P D A G Y Q Q Q Y N P Q G G
235 TATCAACAGTACAATCCTCAAGGCGGTTATCAGCAGCAATTCAATCCACAAGGTGGCCGTGGAAATTACAAAAACTTC
79 Y Q Q Y N P Q G G Y Q Q Q F N P Q G G R G N Y K N F
313 AACTACAATAACAATTTGCAAGGATATCAAGCTGGTTTCCAACCACAGTCTCAAGGTATGTCTTTGAACGACTTTCAA
105 N Y N N N L Q G Y Q A G F Q P Q S Q G M S L N D F Q
391 AAGCAACAAAAGCAGGCCGCTCCCAAACCAAAGAAGACTTTGAAGCTTGTCTCCAGTTCCGGTATCAAGTTGGCCAAT
131 K Q Q K Q A A P K P K K T L K L V S S S G I K L A N
469 GCTACCAAGAAGGTTGGCACAAAACCTGCCGAATCTGATAAGAAAGAGGAAGAGAAGTCTGCTGAAACCAAAGAACCA
157 A T K K V G T K P A E S D K K E E E K S A E T K E P
547 ACTAAAGAGCCAACAAAGGTCGAAGAACCAGTTAAAAAGGAGGAGAAACCAGTCCAGACTGAAGAAAAGACGGAGGAA
183 T K E P T K V E E P V K K E E K P V Q T E E K T E E
625 AAATCGGAACTTCCAAAGGTAGAAGACCTTAAAATCTCTGAATCAACACATAATACCAACAATGCCAATGTTACCAGT
209 K S E L P K V E D L K I S E S T H N T N N A N V T S
← Sup35 GR →
703 GCTGATGCCTTGATCAAGGAACAGGAAGAAGAAGTGGATGACGAAGTTGTTAACGATCCGCGGATGGACTCCAAAGA
235 A D A L I K E Q E E E V D D E V V N D P R M D S K E
780 ATCCTTAGCTCCCCCTGGTAGAGACGAAGTCCCTGGCAGTTTGCTTGGCCAAGGGAGGGGGAGCGTAATGGACTTTTA
260 S L A P P G R D E V P G S L L G Q G R G S V M D F Y
858 TAAAAGCCTGAGGGGAGGAGCTACAGTCAAGGTTTCTGCATCTTCGCCCTCAGTGGCTGCTGCTTCTCAGGCAGATTC
286 K S L R G G A T V K V S A S S P S V A A A S Q A D S
936 CAAGCAGCAGAGGATTCTCCTTGATTTCTCGAAAGGCTCCACAAGCAATGTGCAGCAGCGACAGCAGCAGCAGCAGCA
312 K Q Q R I L L D F S K G S T S N V Q Q R Q Q Q Q Q Q
1014 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCAGGCTTATCCAAAGCCGTTTCACTGTCCATGGGGCT
338 Q Q Q Q Q Q Q Q Q Q Q Q Q P G L S K A V S L S M G L
1092 GTATATGGGAGAGACAGAAACAAAAGTGATGGGGAATGACTTGGGCTACCCACAGCAGGGCCAACTTGGCCTTTCCTC
364 Y M G E T E T K V M G N D L G Y P Q Q G Q L G L S S
1170 TGGGGAAACAGACTTTCGGCTTCTGGAAGAAAGCATTGCAAACCTCAATAGGTCGACCAGCGTTCCAGAGAACCCCAA
390 G E T D F R L L E E S I A N L N R S T S V P E N P K
1248 GAGTTCAACGTCTGCAACTGGGTGTGCTACCCCGACAGAGAAGGAGTTTCCCAAAACTCACTCGGATGCATCTTCAGA
416 S S T S A T G C A T P T E K E F P K T H S D A S S E
1326 ACAGCAAAATCGAAAAAGCCAGACCGGCACCAACGGAGGCAGTGTGAAATTGTATCCCACAGACCAAAGCACCTTTGA
442 Q Q N R K S Q T G T N G G S V K L Y P T D Q S T F D
1404 CCTCTTGAAGGATTTGGAGTTTTCCGCTGGGTCCCCAAGTAAAGACACAAACGAGAGTCCCTGGAGATCAGATCTGTT
468 L L K D L E F S A G S P S K D T N E S P W R S D L L
1482 GATAGATGAAAACTTGCTTTCTCCTTTGGCGGGAGAAGATGATCCATTCCTTCTCGAAGGGAACACGAATGAGGATTG
494 I D E N L L S P L A G E D D P F L L E G N T N E D C
1560 TAAGCCTCTTATTTTACCGGACACTAAACCTAAAATTAAGGATACTGGAGATACAATCTTATCAAGTCCCAGCAGTGT
520 K P L I L P D T K P K I K D T G D T I L S S P S S V
1638 GGCACTACCCCAAGTGAAAACAGAAAAAGATGATTTCATTGAACTTTGCACCCCCGGGGTAATTAAGCAAGAGAAACT
546 A L P Q V K T E K D D F I E L C T P G V I K Q E K L
1716 GGGCCCAGTTTATTGTCAGGCAAGCTTTTCTGGGACAAATATAATTGGTAATAAAATGTCTGCCATTTCTGTTCATGG
572 G P V Y C Q A S F S G T N I I G N K M S A I S V H G
1794 TGTGAGTACCTCTGGAGGACAGATGTACCACTATGACATGAATACAGCATCCCTTTCTCAGCAGCAGGATCAGAAGCC
598 V S T S G G Q M Y H Y D M N T A S L S Q Q Q D Q K P
1872 TGTTTTTAATGTCATTCCACCAATTCCTGTTGGTTCTGAAAACTGGAATAGGTGCCAAGGCTCCGGAGAGGACAGCCT
624 V F N V I P P I P V G S E N W N R C Q G S G E D S L
1950 GACTTCCTTGGGGGCTCTGAACTTCCCAGGCCGGTCAGTGTTTTCTAATGGGTACTCAAGCCCTGGAATGAGACCAGA
650 T S L G A L N F P G R S V F S N G Y S S P G M R P D
2028 TGTAAGCTCTCCTCCATCCAGCTCGTCAGCAGCCACGGGACCACCTCCCAAGCTCTGCCTGGTGTGCTCCGATGAAGC
676 V S S P P S S S S A A T G P P P K L C L V C S D E A
2106 TTCAGGATGTCATTACGGGGTGCTGACATGTGGAAGCTGCAAAGTATTCTTTAAAAGAGCAGTGGAAGGACAGCACAA
702 S G C H Y G V L T C G S C K V F F K R A V E G Q H N

FIG. 1A

```
2184 TTACCTTTGTGCTGGAAGAAACGATTGCATCATTGATAAAATTCGAAGGAAAAACTGCCCAGCATGCCGCTATCGGAA
 728▶  Y  L  C  A  G  R  N  D  C  I  I  D  K  I  R  R  K  N  C  P  A  C  R  Y  R  K
2262 ATGTCTTCAGGCTGGAATGAACCTTGAAGCTCGAAAAACAAAGAAAAAAATCAAAGGGATTCAGCAAGCCACTGCAGG
 754▶  C  L  Q  A  G  M  N  L  E  A  R  K  T  K  K  K  I  K  G  I  Q  Q  A  T  A  G
       2341(GR526)
2340 AGTCTCACAAGACACTTCGGAAAATCCTAACAAAACAATAGTTCCTGCAGCATTACCACAGCTCACCCCTACCTTGGT
 780▶  V  S  Q  D  T  S  E  N  P  N  K  T  I  V  P  A  A  L  P  Q  L  T  P  T  L  V
2418 GTCACTGCTGGAGGTGATTGAACCCGAGGTGTTGTATGCAGGATATGATAGCTCTGTTCCAGATTCAGCATGGAGAAT
 806▶  S  L  L  E  V  I  E  P  E  V  L  Y  A  G  Y  D  S  S  V  P  D  S  A  W  R  I
2496 TATGACCACACTCAACATGTTAGGTGGGCGTCAAGTGATTGCAGCAGTGAAATGGGCAAAGGCGATACTAGGCTTGAG
 832▶  M  T  T  L  N  M  L  G  G  R  Q  V  I  A  A  V  K  W  A  K  A  I  L  G  L  R
2574 AAACTTACACCTCGATGACCAAATGACCCTGCTACAGTACTCATGGATGTTTCTCATGGCATTTGCCTTGGGTTGGAG
 858▶  N  L  H  L  D  D  Q  M  T  L  L  Q  Y  S  W  M  F  L  M  A  F  A  L  G  W  R
2652 ATCATACAGACAATCAAGCGGAAACCTGCTCTGCTTTGCTCCTGATCTGATTATTAATGAGCAGAGAATGTCTCTACC
 884▶  S  Y  R  Q  S  S  G  N  L  L  C  F  A  P  D  L  I  I  N  E  Q  R  M  S  L  P
2730 CTGCATGTATGACCAATGTAAACACATGCTGTTTGTCTCCTCTGAATTACAAAGATTGCAGGTATCCTATGAAGAGTA
 910▶  C  M  Y  D  Q  C  K  H  M  L  F  V  S  S  E  L  Q  R  L  Q  V  S  Y  E  E  Y
2808 TCTCTGTATGAAAACCTTACTGCTTCTCTCCTCAGTTCCTAAGGAAGGTCTGAAGAGCCAAGAGTTATTTGATGAGAT
 936▶  L  C  M  K  T  L  L  L  L  S  S  V  P  K  E  G  L  K  S  Q  E  L  F  D  E  I
2886 TCGAATGACTTATATCAAAGAGCTAGGAAAAGCCATCGTCAAAAGGGAAGGGAACTCCAGTCAGAACTGGCAACGGTT
 962▶  R  M  T  Y  I  K  E  L  G  K  A  I  V  K  R  E  G  N  S  S  Q  N  W  Q  R  F
2964 TTACCAACTGACAAAGCTTCTGGACTCCATGCATGAGGTGGTTGAGAATCTCCTTACCTACTGCTTCCAGACATTTTT
 988▶  Y  Q  L  T  K  L  L  D  S  M  H  E  V  V  E  N  L  L  T  Y  C  F  Q  T  F  L
3042 GGATAAGACCATGAGTATTGAATTCCCAGAGATGTTAGCTGAAATCATCACTAATCAGATACCAAAATATTCAAATGG
1014▶  D  K  T  M  S  I  E  F  P  E  M  L  A  E  I  I  T  N  Q  I  P  K  Y  S  N  G
3120 AAATATCAAAAAGCTTCTGTTTCATCAAAAATGA
1040▶  N  I  K  K  L  L  F  H  Q  K  •
```

FIG. 1B

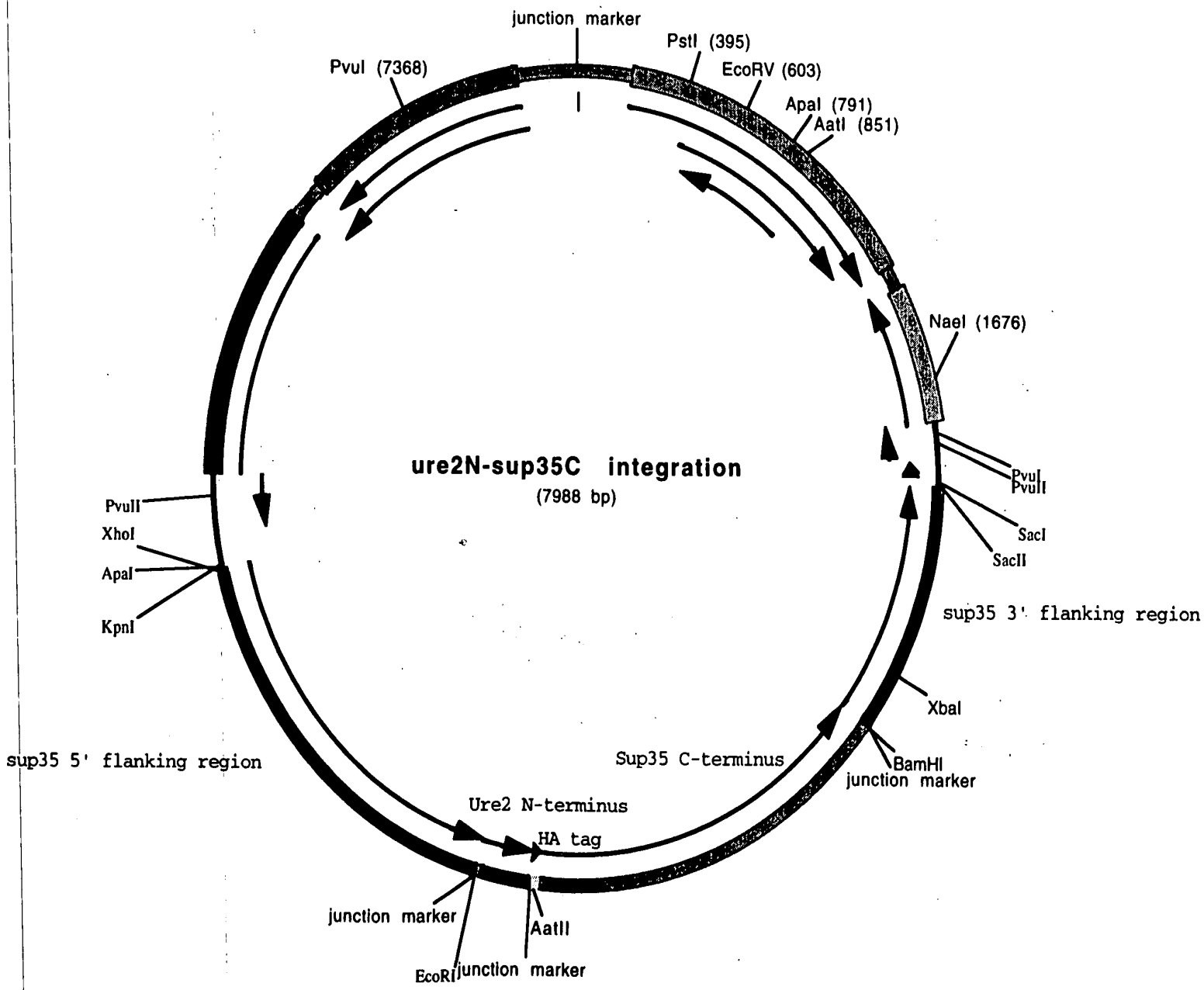

FIG. 2

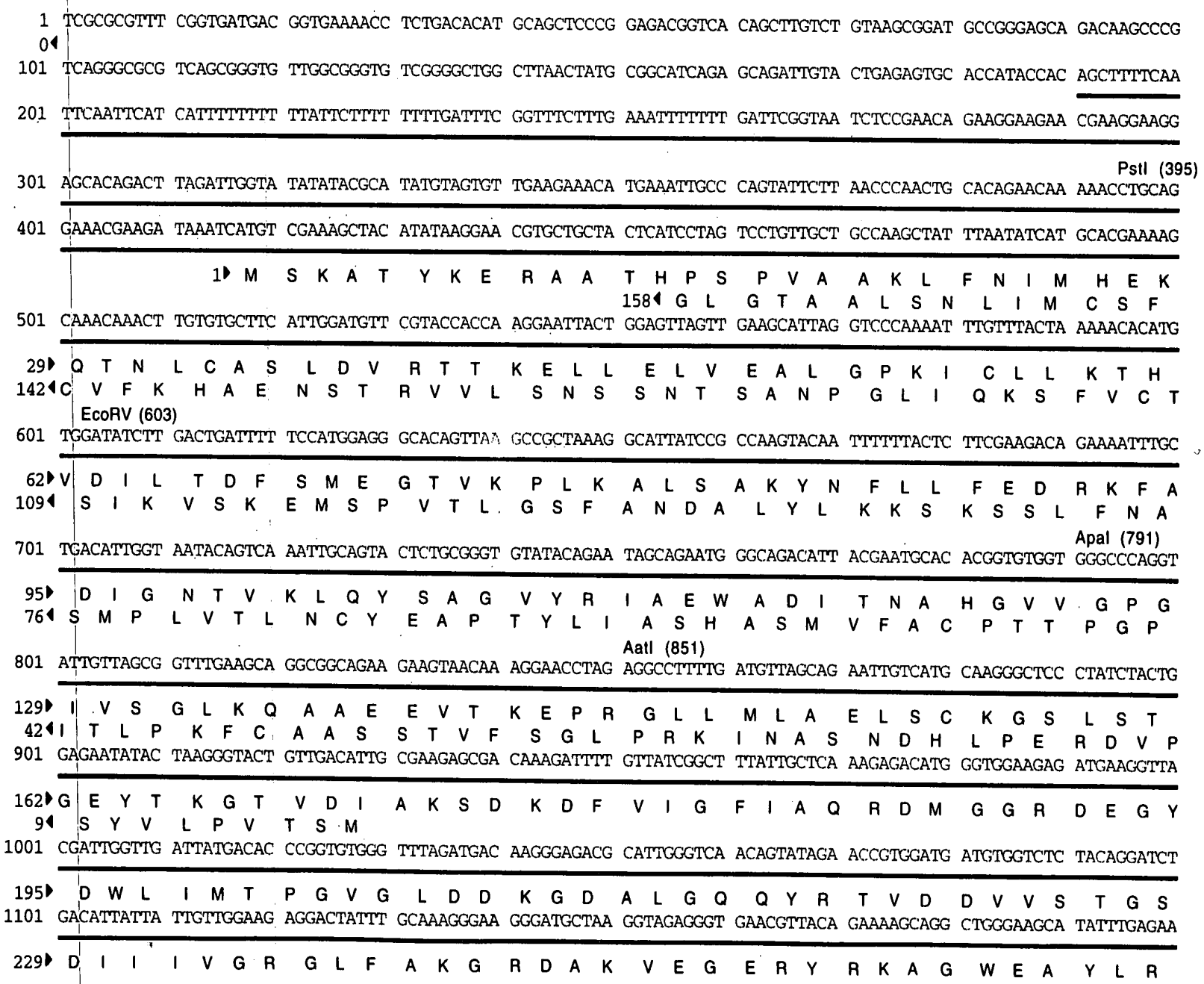
1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
101 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATACCAC AGCTTTTCAA
201 TTCAATTCAT CATTTTTTTT TTATTCTTTT TTTTGATTTC GGTTTCTTTG AAATTTTTTT GATTCGGTAA TCTCCGAACA GAAGGAAGAA CGAAGGAAGG
PstI (395)
301 AGCACAGACT TAGATTGGTA TATATACGCA TATGTAGTGT TGAAGAAACA TGAAATTGCC CAGTATTCTT AACCCAACTG CACAGAACAA AAACCTGCAG
401 GAAACGAAGA TAAATCATGT CGAAAGCTAC ATATAAGGAA CGTGCTGCTA CTCATCCTAG TCCTGTTGCT GCCAAGCTAT TTAATATCAT GCACGAAAAG
1 M S K A T Y K E R A A T H P S P V A A K L F N I M H E K
158 G L G T A A L S N L I M C S F
501 CAAACAAACT TGTGTGCTTC ATTGGATGTT CGTACCACCA AGGAATTACT GGAGTTAGTT GAAGCATTAG GTCCCAAAAT TTGTTTACTA AAAACACATG
29 Q T N L C A S L D V R T T K E L L E L V E A L G P K I C L L K T H
142 C V F K H A E N S T R V V L S N S S N T S A N P G L I Q K S F V C T
EcoRV (603)
601 TGGATATCTT GACTGATTTT TCCATGGAGG GCACAGTTAA GCCGCTAAAG GCATTATCCG CCAAGTACAA TTTTTTACTC TTCGAAGACA GAAAATTTGC
62 V D I L T D F S M E G T V K P L K A L S A K Y N F L L F E D R K F A
109 S I K V S K E M S P V T L G S F A N D A L Y L K K S K S S L F N A
ApaI (791)
701 TGACATTGGT AATACAGTCA AATTGCAGTA CTCTGCGGGT GTATACAGAA TAGCAGAATG GGCAGACATT ACGAATGCAC ACGGTGTGGT GGGCCCAGGT
95 D I G N T V K L Q Y S A G V Y R I A E W A D I T N A H G V V G P G
76 S M P L V T L N C Y E A P T Y L I A S H A S M V F A C P T T P G P
AatII (851)
801 ATTGTTAGCG GTTTGAAGCA GGCGGCAGAA GAAGTAACAA AGGAACCTAG AGGCCTTTTG ATGTTAGCAG AATTGTCATG CAAGGGCTCC CTATCTACTG
129 I V S G L K Q A A E E V T K E P R G L L M L A E L S C K G S L S T
42 I T L P K F C A A S S T V F S G L P R K I N A S N D H L P E R D V P
901 GAGAATATAC TAAGGGTACT GTTGACATTG CGAAGAGCGA CAAAGATTTT GTTATCGGCT TTATTGCTCA AAGAGACATG GGTGGAAGAG ATGAAGGTTA
162 G E Y T K G T V D I A K S D K D F V I G F I A Q R D M G G R D E G Y
9 S Y V L P V T S M
1001 CGATTGGTTG ATTATGACAC CCGGTGTGGG TTTAGATGAC AAGGGAGACG CATTGGGTCA ACAGTATAGA ACCGTGGATG ATGTGGTCTC TACAGGATCT
195 D W L I M T P G V G L D D K G D A L G Q Q Y R T V D D V V S T G S
1101 GACATTATTA TTGTTGGAAG AGGACTATTT GCAAAGGGAA GGGATGCTAA GGTAGAGGGT GAACGTTACA GAAAAGCAGG CTGGGAAGCA TATTTGAGAA
229 D I I I V G R G L F A K G R D A K V E G E R Y R K A G W E A Y L R

FIG. 3A

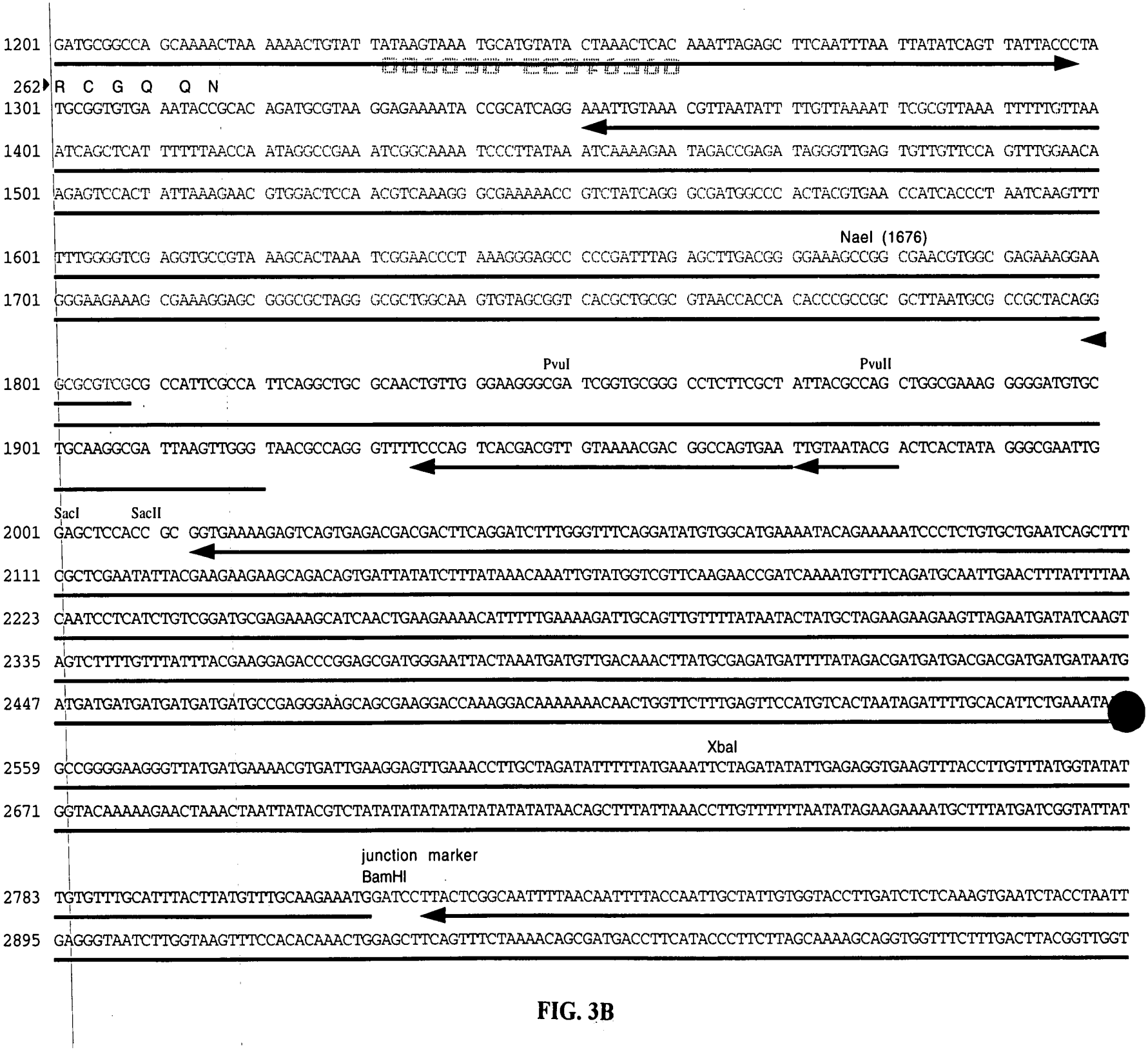
1201 GATGCGGCCA GCAAAACTAA AAAACTGTAT TATAAGTAAA TGCATGTATA CTAAACTCAC AAATTAGAGC TTCAATTTAA TTATATCAGT TATTACCCTA
262 R C G Q Q N
1301 TGCGGTGTGA AATACCGCAC AGATGCGTAA GGAGAAAATA CCGCATCAGG AAATTGTAAA CGTTAATATT TTGTTAAAAT TCGCGTTAAA TTTTTGTTAA
1401 ATCAGCTCAT TTTTTAACCA ATAGGCCGAA ATCGGCAAAA TCCCTTATAA ATCAAAAGAA TAGACCGAGA TAGGGTTGAG TGTTGTTCCA GTTTGGAACA
1501 AGAGTCCACT ATTAAAGAAC GTGGACTCCA ACGTCAAAGG GCGAAAAACC GTCTATCAGG GCGATGGCCC ACTACGTGAA CCATCACCCT AATCAAGTTT
NaeI (1676)
1601 TTTGGGGTCG AGGTGCCGTA AAGCACTAAA TCGGAACCCT AAAGGGAGCC CCCGATTTAG AGCTTGACGG GGAAAGCCGG CGAACGTGGC GAGAAAGGAA
1701 GGGAAGAAAG CGAAAGGAGC GGGCGCTAGG GCGCTGGCAA GTGTAGCGGT CACGCTGCGC GTAACCACCA CACCCGCCGC GCTTAATGCG CCGCTACAGG
PvuI PvuII
1801 GCGCGTCGCG CCATTCGCCA TTCAGGCTGC GCAACTGTTG GGAAGGGCGA TCGGTGCGGG CCTCTTCGCT ATTACGCCAG CTGGCGAAAG GGGGATGTGC
1901 TGCAAGGCGA TTAAGTTGGG TAACGCCAGG GTTTTCCCAG TCACGACGTT GTAAAACGAC GGCCAGTGAA TTGTAATACG ACTCACTATA GGGCGAATTG
SacI SacII
2001 GAGCTCCACC GC GGTGAAAAGAGTCAGTGAGACGACGACTTCAGGATCTTTGGGTTTCAGGATATGTGGCATGAAAATACAGAAAAATCCCTCTGTGCTGAATCAGCTTT
2111 CGCTCGAATATTACGAAGAAGAAGCAGACAGTGATTATATCTTTATAAACAAATTGTATGGTCGTTCAAGAACCGATCAAAATGTTTCAGATGCAATTGAACTTTATTTTAA
2223 CAATCCTCATCTGTCGGATGCGAGAAAGCATCAACTGAAGAAAACATTTTTGAAAAGATTGCAGTTGTTTTATAATACTATGCTAGAAGAAGAAGTTAGAATGATATCAAGT
2335 AGTCTTTTGTTTATTTACGAAGGAGACCCGGAGCGATGGGAATTACTAAATGATGTTGACAAACTTATGCGAGATGATTTTATAGACGATGATGACGACGATGATGATAATG
2447 ATGATGATGATGATGATGATGCCGAGGGAAGCAGCGAAGGACCAAAGGACAAAAAAACAACTGGTTCTTTGAGTTCCATGTCACTAATAGATTTTGCACATTCTGAAATA
XbaI
2559 GCCGGGGAAGGGTTATGATGAAAACGTGATTGAAGGAGTTGAAACCTTGCTAGATATTTTTATGAAATTCTAGATATATTGAGAGGTGAAGTTTACCTTGTTTATGGTATAT
2671 GGTACAAAAAGAACTAAACTAATTATACGTCTATATATATATATATATATATAACAGCTTTATTAAACCTTGTTTTTTAATATAGAAGAAAATGCTTTATGATCGGTATTAT
junction marker
BamHI
2783 TGTGTTTGCATTTACTTATGTTTGCAAGAAATGGATCCTTACTCGGCAATTTTAACAATTTTACCAATTGCTATTGTGGTACCTTGATCTCTCAAAGTGAATCTACCTAATT
2895 GAGGGTAATCTTGGTAAGTTTCCACACAAACTGGAGCTTCAGTTTCTAAAACAGCGATGACCTTCATACCCTTCTTAGCAAAAGCAGGTGGTTTCTTTGACTTACGGTTGGT

FIG. 3B

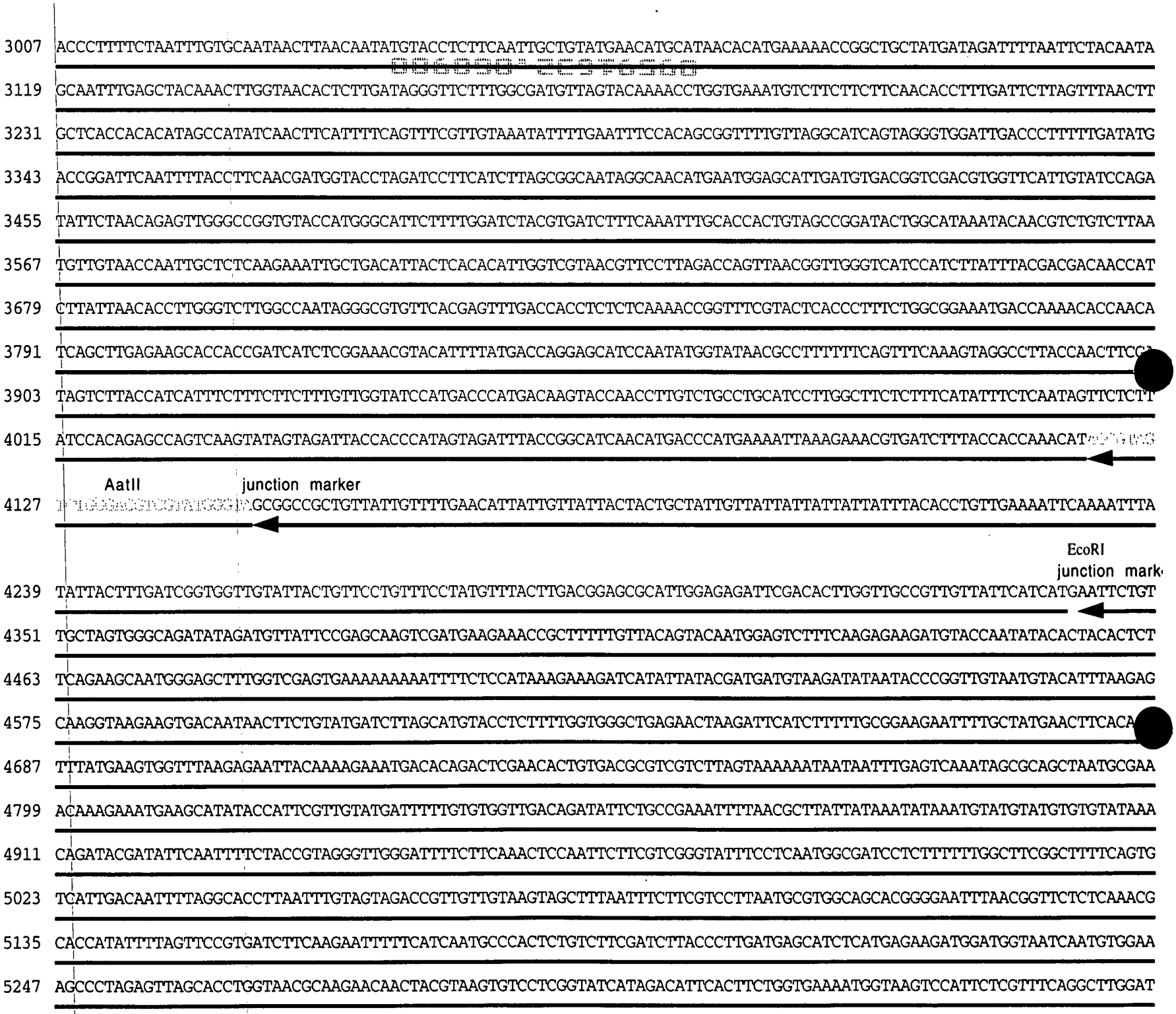
3007 ACCCTTTTCTAATTTGTGCAATAACTTAACAATATGTACCTCTTCAATTGCTGTATGAACATGCATAACACATGAAAAACCGGCTGCTATGATAGATTTTAATTCTACAATA
3119 GCAATTTGAGCTACAAACTTGGTAACACTCTTGATAGGGTTCTTTGGCGATGTTAGTACAAAACCTGGTGAAATGTCTTCTTCTTCAACACCTTTGATTCTTAGTTTAACTT
3231 GCTCACCACACATAGCCATATCAACTTCATTTTCAGTTTCGTTGTAAATATTTTGAATTTCCACAGCGGTTTTGTTAGGCATCAGTAGGGTGGATTGACCCTTTTTGATATG
3343 ACCGGATTCAATTTTACCTTCAACGATGGTACCTAGATCCTTCATCTTAGCGGCAATAGGCAACATGAATGGAGCATTGATGTGACGGTCGACGTGGTTCATTGTATCCAGA
3455 TATTCTAACAGAGTTGGGCCGGTGTACCATGGGCATTCTTTTGGATCTACGTGATCTTTCAAATTTGCACCACTGTAGCCGGATACTGGCATAAATACAACGTCTGTCTTAA
3567 TGTTGTAACCAATTGCTCTCAAGAAATTGCTGACATTACTCACACATTGGTCGTAACGTTCCTTAGACCAGTTAACGGTTGGGTCATCCATCTTATTTACGACGACAACCAT
3679 CTTATTAACACCTTGGGTCTTGGCCAATAGGGCGTGTTCACGAGTTTGACCACCTCTCTCAAAACCGGTTTCGTACTCACCCTTTCTGGCGGAAATGACCAAAACACCAACA
3791 TCAGCTTGAGAAGCACCACCGATCATCTCGGAAACGTACATTTTATGACCAGGAGCATCCAATATGGTATAACGCCTTTTTTCAGTTTCAAAGTAGGCCTTACCAACTTCG
3903 TAGTCTTACCATCATTTCTTTCTTCTTTGTTGGTATCCATGACCCATGACAAGTACCAACCTTGTCTGCCTGCATCCTTGGCTTCTCTTTCATATTTCTCAATAGTTCTCTT
4015 ATCCACAGAGCCAGTCAAGTATAGTAGATTACCACCCATAGTAGATTTACCGGCATCAACATGACCCATGAAAATTAAAGAAACGTGATCTTTACCACCAAACAT
AatII
junction marker
4127 GCGGCCGCTGTTATTGTTTTGAACATTATTGTTATTACTACTGCTATTGTTATTATTATTATTATTTTACACCTGTTGAAAATTCAAAATTTA
EcoRI
junction mark
4239 TATTACTTTGATCGGTGGTTGTATTACTGTTCCTGTTTCCTATGTTTACTTGACGGAGCGCATTGGAGAGATTCGACACTTGGTTGCCGTTGTTATTCATCATGAATTCTGT
4351 TGCTAGTGGGCAGATATAGATGTTATTCCGAGCAAGTCGATGAAGAAACCGCTTTTTGTTACAGTACAATGGAGTCTTTCAAGAGAAGATGTACCAATATACACTACACTCT
4463 TCAGAAGCAATGGGAGCTTTGGTCGAGTGAAAAAAAAATTTTCTCCATAAAGAAAGATCATATTATACGATGATGTAAGATATAATACCCGGTTGTAATGTACATTTAAGAG
4575 CAAGGTAAGAAGTGACAATAACTTCTGTATGATCTTAGCATGTACCTCTTTTGGTGGGCTGAGAACTAAGATTCATCTTTTTGCGGAAGAATTTTGCTATGAACTTCACA
4687 TTTATGAAGTGGTTTAAGAGAATTACAAAAGAAATGACACAGACTCGAACACTGTGACGCGTCGTCTTAGTAAAAAATAATAATTTGAGTCAAATAGCGCAGCTAATGCGAA
4799 ACAAAGAAATGAAGCATATACCATTCGTTGTATGATTTTTGTGTGGTTGACAGATATTCTGCCGAAATTTTAACGCTTATTATAAATATAAATGTATGTATGTGTGTATAAA
4911 CAGATACGATATTCAATTTTCTACCGTAGGGTTGGGATTTTCTTCAAACTCCAATTCTTCGTCGGGTATTTCCTCAATGGCGATCCTCTTTTTTGGCTTCGGCTTTTCAGTG
5023 TCATTGACAATTTTAGGCACCTTAATTTGTAGTAGACCGTTGTTGTAAGTAGCTTTAATTTCTTCGTCCTTAATGCGTGGCAGCACGGGGAATTTAACGGTTCTCTCAAACG
5135 CACCATATTTTAGTTCCGTGATCTTCAAGAATTTTTCATCAATGCCCACTCTGTCTTCGATCTTACCCTTGATGAGCATCTCATGAGAAGATGGATGGTAATCAATGTGGAA
5247 AGCCCTAGAGTTAGCACCTGGTAACGCAAGAACAACTACGTAAGTGTCCTCGGTATCATAGACATTCACTTCTGGTGAAAATGGTAAGTCCATTCTCGTTTCAGGCTTGGAT

FIG. 3C

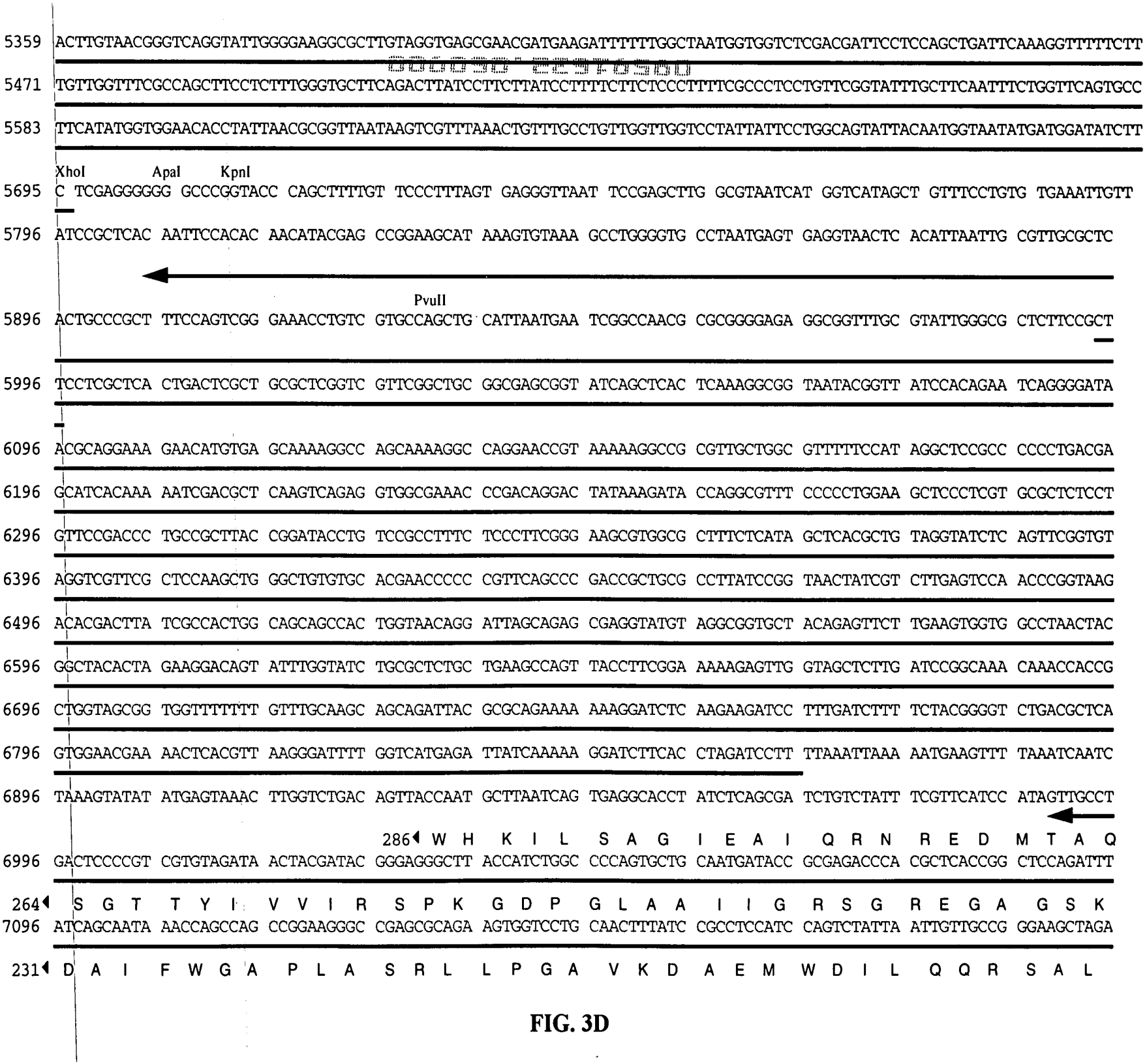
5359 ACTTGTAACGGGTCAGGTATTGGGGAAGGCGCTTGTAGGTGAGCGAACGATGAAGATTTTTTGGCTAATGGTGGTCTCGACGATTCCTCCAGCTGATTCAAAGGTTTTTCTT
5471 TGTTGGTTTCGCCAGCTTCCTCTTTGGGTGCTTCAGACTTATCCTTCTTATCCTTTTCTTCTCCCTTTTCGCCCTCCTGTTCGGTATTTGCTTCAATTTCTGGTTCAGTGCC
5583 TTCATATGGTGGAACACCTATTAACGCGGTTAATAAGTCGTTTAAACTGTTTGCCTGTTGGTTGGTCCTATTATTCCTGGCAGTATTACAATGGTAATATGATGGATATCTT
XhoI ApaI KpnI
5695 C TCGAGGGGGG GCCCGGTACC CAGCTTTTGT TCCCTTTAGT GAGGGTTAAT TCCGAGCTTG GCGTAATCAT GGTCATAGCT GTTTCCTGTG TGAAATTGTT
5796 ATCCGCTCAC AATTCCACAC AACATACGAG CCGGAAGCAT AAAGTGTAAA GCCTGGGGTG CCTAATGAGT GAGGTAACTC ACATTAATTG CGTTGCGCTC
PvuII
5896 ACTGCCCGCT TTCCAGTCGG GAAACCTGTC GTGCCAGCTG CATTAATGAA TCGGCCAACG CGCGGGGAGA GGCGGTTTGC GTATTGGGCG CTCTTCCGCT
5996 TCCTCGCTCA CTGACTCGCT GCGCTCGGTC GTTCGGCTGC GGCGAGCGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT ATCCACAGAA TCAGGGGATA
6096 ACGCAGGAAA GAACATGTGA GCAAAAGGCC AGCAAAAGGC CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC CCCCTGACGA
6196 GCATCACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAAC CCGACAGGAC TATAAAGATA CCAGGCGTTT CCCCCTGGAA GCTCCCTCGT GCGCTCTCCT
6296 GTTCCGACCC TGCCGCTTAC CGGATACCTG TCCGCCTTTC TCCCTTCGGG AAGCGTGGCG CTTTCTCATA GCTCACGCTG TAGGTATCTC AGTTCGGTGT
6396 AGGTCGTTCG CTCCAAGCTG GGCTGTGTGC ACGAACCCCC CGTTCAGCCC GACCGCTGCG CCTTATCCGG TAACTATCGT CTTGAGTCCA ACCCGGTAAG
6496 ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG GCCTAACTAC
6596 GGCTACACTA GAAGGACAGT ATTTGGTATC TGCGCTCTGC TGAAGCCAGT TACCTTCGGA AAAAGAGTTG GTAGCTCTTG ATCCGGCAAA CAAACCACCG
6696 CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC AGCAGATTAC GCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT TCTACGGGGT CTGACGCTCA
6796 GTGGAACGAA AACTCACGTT AAGGGATTTT GGTCATGAGA TTATCAAAAA GGATCTTCAC CTAGATCCTT TTAAATTAAA AATGAAGTTT TAAATCAATC
6896 TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT TCGTTCATCC ATAGTTGCCT
286 W H K I L S A G I E A I Q R N R E D M T A Q
6996 GACTCCCCGT CGTGTAGATA ACTACGATAC GGGAGGGCTT ACCATCTGGC CCCAGTGCTG CAATGATACC GCGAGACCCA CGCTCACCGG CTCCAGATTT
264 S G T T Y I V V I R S P K G D P G L A A I I G R S G R E G A G S K
7096 ATCAGCAATA AACCAGCCAG CCGGAAGGGC CGAGCGCAGA AGTGGTCCTG CAACTTTATC CGCCTCCATC CAGTCTATTA ATTGTTGCCG GGAAGCTAGA
231 D A I F W G A P L A S R L L P G A V K D A E M W D I L Q Q R S A L

FIG. 3D

```
7196 GTAAGTAGTT CGCCAGTTAA TAGTTTGCGC AACGTTGTTG CCATTGCTAC AGGCATCGTG GTGTCACGCT CGTCGTTTGG TATGGCTTCA TTCAGCTCCG

197 T L L E G T L L K R L T T A M A V P M T T D R E D N P I A E N L E P

                                                                    PvuI (7368)
7296 GTTCCCAACG ATCAAGGCGA GTTACATGAT CCCCCATGTT GTGCAAAAAA GCGGTTAGCT CCTTCGGTCC TCCGATCGTT GTCAGAAGTA AGTTGGCCGC

164  E W R D L R T V H D G M N H L F A T L E K P G G I T T L L L N A A
7396 AGTGTTATCA CTCATGGTTA TGGCAGCACT GCATAATTCT CTTACTGTCA TGCCATCCGT AAGATGCTTT TCTGTGACTG GTGAGTACTC AACCAAGTCA

131 T N D S M T I A A S C L E R V T M G D T L H K E T V P S Y E V L D
7496 TTCTGAGAAT AGTGTATGCG GCGACCGAGT TGCTCTTGCC CGGCGTCAAT ACGGGATAAT ACCGCGCCAC ATAGCAGAAC TTTAAAAGTG CTCATCATTG

97 N Q S Y H I R R G L Q E Q G A D I R S L V A G C L L V K F T S M M P
7596 GAAAACGTTC TTCGGGGCGA AAACTCTCAA GGATCTTACC GCTGTTGAGA TCCAGTTCGA TGTAACCCAC TCGTGCACCC AACTGATCTT CAGCATCTTT

64  F R E E P R F S E L I K G S N L D L E I Y G V R A G L Q D E A D K
7696 TACTTTCACC AGCGTTTCTG GGTGAGCAAA AACAGGAAGG CAAAATGCCG CAAAAAAGGG AATAAGGGCG ACACGGAAAT GTTGAATACT CATACTCTTC

31 V K V L T E P H A F V P L C F A A F F P I L A V R F H Q I S M
7796 CTTTTTCAAT ATTATTGAAG CATTTATCAG GGTTATTGTC TCATGAGCGG ATACATATTT GAATGTATTT AGAAAAATAA ACAAATAGGG GTTCCGCGCA

                                                                    junction marker
7896 CATTTCCCCG AAAAGTGCCA CCTGACGTCT AAGAAACCAT TATTATCATG ACATTAACCT ATAAAAATAG GCGTATCACG AGGCCCTTTC GTC
```

FIG. 3E

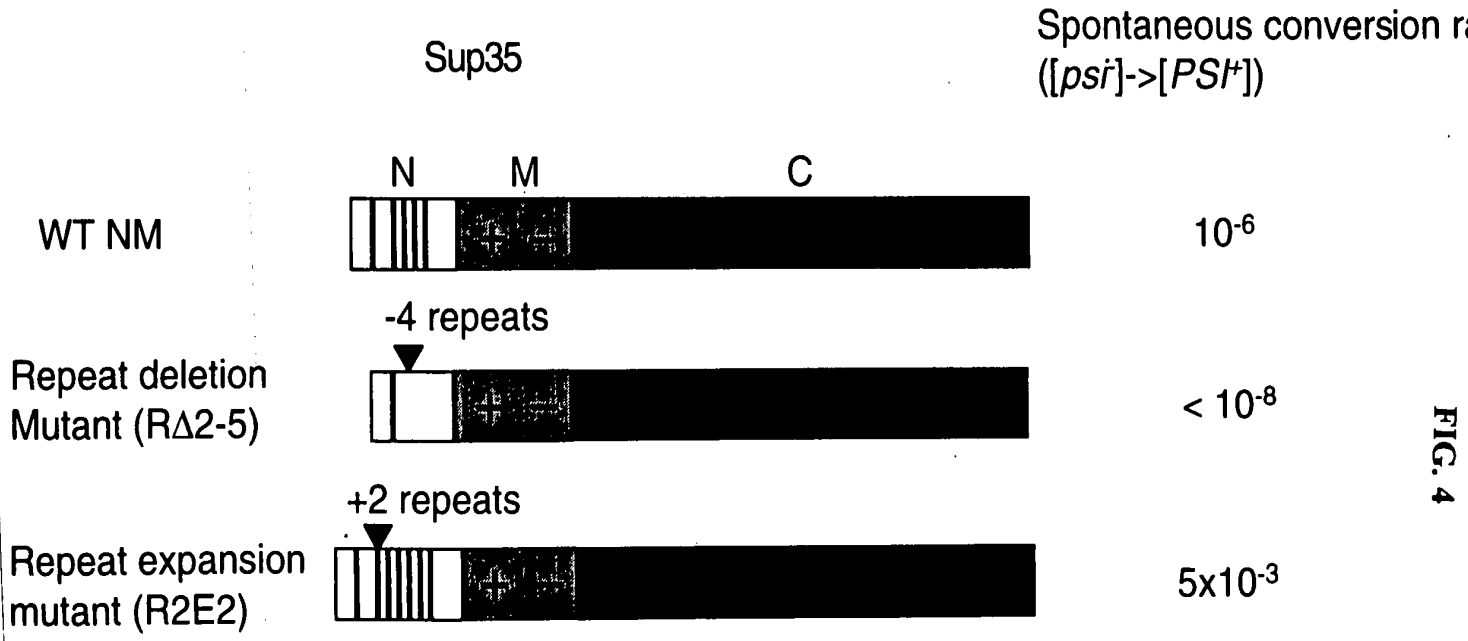

FIG. 4

Spontaneous conversion of Sup35 repeat mutants